ON QUATERNIONS AND THEIR GENERALIZATION AND THE HISTORY OF THE EIGHT SQUARE THEOREM.

By L. E. Dickson.

1. **Objects of the paper.** We shall present the history of the generalizations to four and eight squares of the familiar formula

$$(1)\quad (a^2 + b^2)(\alpha^2 + \beta^2) = r^2 + s^2, \qquad r = a\alpha - b\beta, \qquad s = a\beta + b\alpha,$$

and an elementary exposition of Hurwitz's proof that such a formula holds only for 2, 4 or 8 squares. For these three cases we shall show that the formula admits of a simple interpretation concerning the norms of numbers which are ordinary complex numbers, quaternions or numbers of Cayley's algebra with 8 units. No knowledge of quaternions or the latter algebra will be presupposed, but their more fundamental algebraic properties will be developed in detail.

A clear exposition will first be given (§§ 1–5) of the main results of our subject. This will be followed (§§ 6–28) by an account of its history, which is believed to omit no paper on the eight square theorem and its generalization

2. **Ordinary complex numbers.** Let a and b be any real numbers. Then the complex number $a + bi$ is said to have the *norm* $a^2 + b^2$. Formula (1) evidently expresses the property that the norm of the product $r + si$ of the complex numbers $a + bi$ and $\alpha + \beta i$ equals the product of their norms.

To prepare the way for our introduction to quaternions and Cayley's algebra, we shall present briefly W. R. Hamilton's definition of complex numbers by means of couples of real numbers. Two couples (a, b) and (α, β) are called equal if and only if $a = \alpha$, $b = \beta$. Addition and multiplication are defined by

$$(a, b) + (\alpha, \beta) = (a + \alpha, b + \beta), \qquad (a, b)(\alpha, \beta) = (r, s),$$

where r and s are given by (1). If m is any real number, we define $m(a, b)$ and $(a, b)m$ to be (ma, mb). Writing 1 for $(1, 0)$ and i for $(0, 1)$, we have

$$(a, b) = (a, 0) + (0, b) = a(1, 0) + b(0, 1) = a + bi.$$

The previous definition of addition and multiplication of couples gives

$$(a + bi) + (\alpha + \beta i) = a + \alpha + (b + \beta)i, \qquad (a + bi)(\alpha + \beta i) = r + si.$$

3. **Quaternions.** Consider quadruples (a, b, c, d) of real or complex numbers a, b, c, d. Define addition and multiplication by

$$(a, b, c, d) + (\alpha, \beta, \gamma, \delta) = (a + \alpha, b + \beta, c + \gamma, d + \delta),$$

$$(a, b, c, d) \times (\alpha, \beta, \gamma, \delta) = (A, B, C, D),$$

where

$$\begin{cases} A = a\alpha - b\beta - c\gamma - d\delta, & B = a\beta + b\alpha + c\delta - d\gamma, \\ C = a\gamma - b\delta + c\alpha + d\beta, & D = a\delta + b\gamma - c\beta + d\alpha. \end{cases} \tag{2}$$

No attempt will be made here to explain why we select these values for $A, \cdots, D$; it is not our purpose to explain how quaternions were discovered or how they may be made to enter naturally,* as we aim merely to give a logical basis for quaternions. Consider the four particular quadruples

$$1 = (1, 0, 0, 0), \quad i = (0, 1, 0, 0), \quad j = (0, 0, 1, 0), \quad k = (0, 0, 0, 1),$$

called the *units*. Define $m(a, b, c, d)$ or $(a, b, c, d)m$ to be (ma, mb, mc, md), where m is any complex number. Then

$$(a, b, c, d) = (a, 0, 0, 0) + \cdots + (0, 0, 0, d) = a + bi + cj + dk,$$

$$\begin{aligned} i^2 = j^2 = k^2 = -1, \quad & ij = k, \quad ji = -k, \\ & jk = i, \quad kj = -i, \quad ki = j, \quad ik = -j. \end{aligned} \tag{3}$$

Henceforth we discard the quadruple notation and employ

$$q = a + bi + cj + dk, \qquad Q = \alpha + \beta i + \gamma j + \delta k,$$

called quaternions. In view of our earlier definitions, their sum is $a + \alpha + (b + \beta)i + \cdots$ and their product is $A + Bi + Cj + Dk$, where $A, \cdots, D$ have the values (2). This product may be found by performing the multiplication as in formal algebra, care being taken not to permute two factors i, j, k, and then simplifying the result by use of (3). For example, $(i + 2j)(j + k) = k - j - 2 + 2i$. Note that, while multiplication is not commutative, it is associative since $(ij)k = -1 = i(jk)$, etc.

The quaternion $q' = a - bi - cj - dk$ is called the *conjugate* to q. We readily verify that $qq' = q'q = a^2 + b^2 + c^2 + d^2$, which is called the *norm* $N(q)$ of q. For the moment, let a, b, c, d be real numbers, so that q is a real quaternion; if $q \neq 0$, then $N(q) \neq 0$, and q has the inverse $q^{-1} = q'/N(q)$. Thus, if $q \neq 0$, $qQ = q_1$ has the unique solution $Q = q^{-1}q_1$, and $Qq = q_1$ has the unique solution $Q = q_1q^{-1}$, so that both right-hand

* This topic is presented in an elementary manner in Dickson's Linear Algebras, Cambridge University Tract No. 16, pp. 9–12, and from another standpoint in his article "On the relation between linear algebras and continuous groups," Bull. Amer. Math. Soc., 22, 1915, 53–61.

and left-hand division are always uniquely possible if the divisor is a real quaternion not zero.

The conjugate of qQ equals the product $Q'q'$ of the conjugates of the factors taken in reverse order, as shown by interchanging the Roman and Greek letters in the sums (2) and afterwards changing the signs of $b, c, d, \beta, \gamma, \delta$.

The norm of qQ is $qQ \cdot Q'q'$ by definition. By the associative law, this may be written $q(QQ')q'$. Since QQ' is an ordinary number, it is commutative with q' in view of our earlier definition of $m(a, b, c, d)$ and $(a, b, c, d)m$. The result is now the product of the norms qq' and QQ' of q and Q. Hence the norm of a product of two quaternions equals the product of their norms, *i. e.*,

$$(4)\quad \begin{aligned}&(a^2 + b^2 + c^2 + d^2)(\alpha^2 + \beta^2 + \gamma^2 + \delta^2)\\ &\qquad = A^2 + B^2 + C^2 + D^2 \qquad (A, \cdots, D \text{ as in } (2)).\end{aligned}$$

Much earlier than Hamilton's invention of quaternions in 1843, Euler* discovered formula (4) while investigating the elegant theorem that every positive integer is a sum of four integral squares, the theorem following from (4) if proved for every prime number; he also used (4) in his later paper on orthogonal substitutions.

4. **Cayley's algebra.** A. Cayley† defined an algebra with the 8 units $1, i_1, \cdots, i_7$, such that $i_1^2 = -1, \cdots, i_7^2 = -1$,

$$i_1 i_2 = i_3 = -i_2 i_1, \quad i_2 i_3 = i_1 = -i_3 i_2, \quad i_3 i_1 = i_2 = -i_1 i_3,$$

and six similar sets of six relations with 1, 2, 3 replaced by 1, 4, 5; 6, 2, 4; 6, 5, 3; 7, 2, 5; 7, 3, 4; 1, 7, 6; respectively. Then

$$(x_0 + x_1 i_1 + \cdots + x_7 i_7)(x_0' + x_1' i_1 + \cdots + x_7' i_7) = A_0 + A_1 i_1 + \cdots + A_7 i_7,$$

where, if we employ the abbreviations $jk = x_j x_k' - x_k x_j'$, $\overline{0j} = x_0 x_j' + x_j x_0'$,

$$\begin{aligned}
A_0 &= x_0 x_0' - x_1 x_1' - \cdots - x_7 x_7', & A_1 &= 23 + 45 + 76 + \overline{01},\\
A_2 &= 31 + 46 + 57 + \overline{02}, & A_3 &= 12 + 65 + 47 + \overline{03},\\
A_4 &= 51 + 62 + 73 + \overline{04}, & A_5 &= 14 + 36 + 72 + \overline{05},\\
A_6 &= 24 + 53 + 17 + \overline{06}, & A_7 &= 25 + 34 + 61 + \overline{07}.
\end{aligned}$$

He called Σx_i^2 the modulus (norm) of $x_0 + \cdots + x_7 i_7$ and stated that the

* Corresp. Math. Phys. (ed., P. H. Fuss), I, 1843, 452, letter to Goldbach, May 4, 1748. Novi Comm. Acad. Petrop., 5, 1754–5, 3; 15, 1770, 75; Comm. Arith. Coll., I, 230, 427.

† Phil. Mag., London, (3), 26, 1845, 210 [30, 1847, 257–8]; Coll. Math. Papers, I, 127 [301]. In A_4 his 87 is a misprint for 73.

norm of a product equals the product of the norms of the two factors:

$$\left(\sum_{i=0}^{7} x_i^2\right)\left(\sum_{i=0}^{7} x_i'^2\right) = \sum_{i=0}^{7} A_i^2. \tag{5}$$

The last result, as well as another important property of the algebra, can be proved without computation by representing the algebra as a quasi-binary algebra.* Since $1, i_1, i_2, i_3$ satisfy the relations (3) for the quaternion units, we may replace them by $1, i, j, k$. Then the remaining four units are $e = i_4, ie = i_5, je = i_6, ke = i_7$. Hence every number of the algebra is of the form $q + Qe$, where q and Q are linear functions of $1, i, j, k$ and hence are quaternions. It can be verified that Cayley's 49 relations, giving the product of two equal or distinct units $i_1, \cdots, i_7$, are together equivalent to the single formula

$$(q + Qe)(r + Re) = qr - R'Q + (Rq + Qr')e, \tag{6}$$

where r' and R' are the quaternions conjugate to r and R. The reader need not verify the equivalence stated, but may take (6) as the rule of multiplication for the numbers of the algebra to be considered henceforth, since Cayley's algebra has been introduced here merely for historical background and will not be further employed in his form.

Define the norm of $q + Qe$ to be $qq' + QQ'$, which is a sum of 8 squares. Taking $r = q'$, $R = -Q$, in (6), we get

$$(q + Qe)(q' - Qe) = qq' + QQ',$$

so that the norm of $q + Qe$ is its product by its *conjugate* $q' - Qe$. Since multiplication does not here obey the associative law, we cannot conclude at once, as we did for quaternions in §3, that the norm of a product equals the product of the norms of the two factors. However, we obtain a short proof by use of a device. Express the right member of (6) in the form $t + Te$ by setting

$$t = qr - R'Q, \qquad T = Rq + Qr'. \tag{7}$$

Its norm $tt' + TT'$ is seen, by direct multiplication and use of the fact that the norm of qr is $r'q'$, to equal $\alpha - \beta + \gamma$, where

$$\alpha = RqrQ' + Qr'q'R', \qquad \beta = qrQ'R + R'Qr'q',$$

$$\gamma = qrr'q' + R'QQ'R + Rqq'R' + Qr'rQ' = (qq' + QQ')(rr' + RR').$$

The last equality is a consequence of the fact that rr' is an ordinary number and hence can be interchanged with q', etc. Our device occurs in the proof that $\alpha = \beta$. Note that the conjugate of the first term of α

* Dickson, Trans. Amer. Math. Soc., 13, 1912, 72; Linear Algebras, 1914, 15.

equals the second term of α, so that α is an ordinary number and hence is commutative with every quaternion. Hence $\alpha = R'\alpha R \div RR'$, which is seen to equal β. In the excluded case $R = 0$, evidently $\alpha = \beta = 0$. Hence the norm of the product (6) equals the product of the norms of the factors. Thus we can write down an 8 square formula of type (5).

Moreover, both right-hand and left-hand division except by zero is always possible and unique in our algebra composed of the numbers $q + Qe$, provided we restrict q and Q to be real quaternions. Of the two types of division consider that in which the second factor $r + Re$ and the product $t + Te$ are given, while the first factor $q + Qe$ is to be found. Thus we seek to solve equations (7) for q and Q. Multiply the second equation (7) by r on the right and replace qr by its value from the first equation; we get

$$(rr' + RR')Q = Tr - Rt.$$

Again, multiply the first equation by r' on the right and eliminate Qr'; thus

$$(rr' + RR')q = tr' + R'T.$$

Since $rr' + RR'$ equals the sum of the squares of eight real numbers, it is zero if and only if $r = R = 0$. Similarly, equations (7) can be solved for $r + Re$ unless $q = Q = 0$.

We have now accomplished one of the aims of the paper, having exhibited linear algebras in 2, 4 and 8 units for which the norm of a product equals the product of the norms of the factors (thus giving the 2, 4 and 8 square theorems), and such that, if the coördinates of the numbers of the algebra be restricted to be real numbers, both right-hand and left-hand division except by zero are possible and unique. While the three algebras have in common these two fundamental properties, they differ in other respects. For complex numbers multiplication is both commutative and associative, for quaternions it is associative but not commutative, for Cayley's algebra of 8 units it is neither commutative nor associative. What additional properties must be given up to obtain a similar linear algebra in more than 8 units? We shall prove in §5 that there exists no linear algebra in more than 8 units for which the norm is a sum of squares and the norm of a product equals the product of the norms of the factors.

5. **Hurwitz's Theorem.*** We seek the values of n for which there exists

* Göttingen Nachrichten, 1898, 309–316. Since experience shows that graduate students fail to follow various steps merely outlined by Hurwitz, we shall here give the proof in detailed, amplified form. As we shall employ (a_{ij}) to denote a matrix and not a linear transformation, we must invert the order of factors in his products.

an identity (as to the x's and y's) of the form

$$(8)\qquad (x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2) = z_1^2 + \cdots + z_n^2,$$

where $z_1, \cdots, z_n$ are linear in $x_1, \cdots, x_n$ and also in $y_1, \cdots, y_n$. Let

$$(9)\qquad z_i = a_{i1}y_1 + \cdots + a_{in}y_n \qquad (i = 1, \cdots, n),$$

where the a_{ij} are linear functions of $x_1, \cdots, x_n$. We employ the matrices

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \qquad A' = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix},$$

where A' is derived from A by the interchange of its rows and columns, and is called the conjugate (or transposed) of A. In case the diagonal elements a_{ii} all equal a and the elements not in the diagonal are all zero, we shall write aI for A, where I is the unit (or identity) matrix and has the property that $IB = BI = B$ for every matrix B of n rows and n columns. The quadratic form

$$\sum_{i,\, j=1}^{n} b_{ij}z_iz_j = b_{11}z_1^2 + 2b_{12}z_1z_2 + b_{22}z_2^2 + \cdots \qquad (b_{ij} = b_{ji})$$

is said to have the matrix $B = (b_{ij})$, whose ith row is $b_{i1}, b_{i2}, \cdots, b_{in}$. If we replace the variables $z_1, \cdots, z_n$ by the expressions (9), we evidently obtain a new quadratic form in the variables $y_1, \cdots, y_n$; its matrix is known* to equal $A'BA$. In particular, let the quadratic form be $z_1^2 + \cdots + z_n^2$, whose matrix is $B = I$; then the quadratic form derived by replacing $z_1, \cdots, z_n$ by the expressions (9) has the matrix $A'A$, a fact which can be verified at once without making use of the standard theorem just quoted. Now we desire that the resulting quadratic form in $y_1, \cdots, y_n$ shall be the left member of (8), whose matrix is aI, where $a = x_1^2 + \cdots + x_n^2$. Hence there exists an identity (8) if and only if there exist n^2 linear functions a_{ij} of $x_1, \cdots, x_n$, whose matrix (a_{ij}) is denoted by A, such that

$$(10)\qquad A'A = (x_1^2 + \cdots + x_n^2)I.$$

Since each element of matrix A is a linear function of $x_1, \cdots, x_n$, and since the sum of several matrices is a matrix whose elements are the sums of the corresponding elements in the matrices added, it follows that $A = x_1A_1 + \cdots + x_nA_n$, where $A_1, \cdots, A_n$ are matrices with constant elements. Thus in $A'A$ the coefficient of x_n^2 is $A_n'A_n$, which equals I by (10). Let $B_i = A_n'A_i$ $(i = 1, \cdots, n - 1)$, whence $A_i = A_nB_i$, A_i'

* Bôcher, Introduction to Higher Algebra, p. 129.

$= B_i' A_n'$, and $A'A$ equals

$$(x_1 B_1' + \cdots + x_{n-1} B_{n-1}' + x_n) A_n' \cdot A_n (x_1 B_1 + \cdots + x_{n-1} B_{n-1} + x_n).$$

Since $A_n' A_n = I$, (10) becomes

$$(11) \quad (x_1 B_1' + \cdots + x_{n-1} B_{n-1}' + x_n)(x_1 B_1 + \cdots + x_{n-1} B_{n-1} + x_n) = (x_1^2 + \cdots + x_n^2) I.$$

Thus $B_i' B_i = I$, $B_i' + B_i = 0$, $B_i' B_k + B_k' B_i = 0$, whence

$$(12) \quad B_i' = -B_i, \quad B_i^2 = -I, \quad B_i B_k = -B_k B_i \qquad (i, k = 1, \cdots, n-1; i \neq k).$$

A matrix $B = (b_{ij})$ is called symmetric if $b_{ji} = b_{ij}$, and skew symmetric if $b_{ji} = -b_{ij}$ for every i, j; thus B is symmetric if and only if $B' = B$, and skew-symmetric if and only if $B' = -B$. The latter condition implies that $b = (-)^n b$ if b is the determinant of the matrix B of n rows and n columns. Thus $b = 0$ if n is odd. By the first two equations (12), B_i is skew-symmetric and its determinant is not zero, so that n is not odd. Hence there exists no identity (8) if n is odd. In what follows, we assume that n is even.

Our next step is to prove that at least half of the matrices

$$(13) \quad I, \quad B_{i_1}, \quad B_{i_1} B_{i_2}, \quad B_{i_1} B_{i_2} B_{i_3}, \quad \cdots, \quad B_1 B_2 \cdots B_{n-1} \qquad (i_1 < n, \; i_1 < i_2 < n, \; \cdots)$$

are linearly independent. There are 2^{n-1} such products since any one product either contains B_1 or does not, $\cdots$, and either contains B_{n-1} or does not. Let $G = B_{i_1} \cdots B_{i_r}$ be one of the matrices (13); it is symmetric if $r \equiv 0$ or $3 \pmod 4$, and skew-symmetric if $r \equiv 1$ or $2 \pmod 4$, since by (12)

$$G' = B_{i_r}' \cdots B_{i_1}' = (-1)^r B_{i_r} \cdots B_{i_1} = (-1)^s G,$$

where $s = r + r - 1 + r - 2 + \cdots + 1 = r(r+1)/2$ is even if $r \equiv 0, 3 \pmod 4$, but odd if $r \equiv 1, 2 \pmod 4$. In particular, a product of two distinct B's is skew-symmetric.

Consider the possible linear relations (with constant coefficients not all zero) which hold between the matrices (13). Such a relation $R = 0$ is called irreducible if it is not possible to express R in the form $R = R_1 + R_2$, where $R_1 = 0$ and $R_2 = 0$ represent two linear relations holding between our matrices such that no one of these matrices (13) occurs as a term of both R_1 and R_2. In particular, an irreducible linear relation does not involve both symmetric and skew-symmetric matrices, since it could

then be written in the form $M = S$, where M is the aggregate of the symmetric matrices and S is the aggregate of the skew symmetric matrices, whence $M' = S'$, $M' = M$, $S' = -S$, giving $M = 0$, $S = 0$.

Let $R = 0$ be any irreducible linear relation between the matrices (13). By multiplying R by the product of a constant and a suitably chosen matrix (13), we get a new linear relation $\rho = 0$, one term of which is I and all the remaining terms are products of matrices (13) by constants. Thus if $4B_2B_3$ is one term of R, we use the multiplier $-\frac{1}{4}B_2B_3$. We need also to know that if we multiply the matrices (13) on the left by any one (say M) of them, the products form a permutation of those matrices each prefixed with the factor $+1$ or -1. This is evident when the multiplier is B_1, since the product will contain or lack B_1 according as the multiplicand (13) lacks or contains B_1, in view of $B_1^2 = -I$. If the multiplier is B_2, we first replace $B_1B_2 \cdots$ by $-B_2B_1 \cdots$ and see that the former argument applies. After proving in this manner our statement when the multiplier is any B_i, we see that it holds when the multiplier is any product of the B's. Returning to our new relation $\rho = 0$, we note that it also is irreducible, since by multiplying it by a product of a constant and a suitable matrix (13) we recover our initial relation $R = 0$, which was assumed irreducible. Hence $\rho = 0$ is an irreducible relation

$$I = \Sigma c_{i_1i_2i_3}B_{i_1i_2i_3} + \Sigma d_{i_1i_2i_3i_4}B_{i_1}B_{i_2}B_{i_3}B_{i_4} + \cdots$$

involving exclusively symmetric matrices (13), so that no term contains a single B_i or a product of only two B's. Multiply all the terms of our relation by B_i on the right; we obtain an irreducible relation which therefore involves only skew-symmetric matrices (13), one term being B_i. Since a product of four distinct B's is symmetric, we conclude that $c_{i_1i_2i_3}$ is zero if i is distinct from i_1, i_2, i_3. Since i may have any value $\leqq n-1$, we have $c = 0$ unless $3 = n - 1$. To prove that every $d = 0$, take $i = i_4$; then the coefficient of $-dB_{i_1}B_{i_2}B_{i_3}$ is zero. The method used to prove $c = 0$ applies when the number r of factors B is $\equiv 3 \pmod 4$ and $r < n - 1$, since $r + 1 \equiv 0$. The method used to prove $d = 0$ applies when $r \equiv 0 \pmod 4$, since $r - 1 \equiv 3$. Hence if our relation exists, it has the form

$$I = kB_1B_2 \cdots B_{n-1}.$$

Since each member is a symmetric matrix, $n - 1 \equiv 0$ or $3 \pmod 4$. But n is even. Hence $n \equiv 0 \pmod 4$. As in the discussion of G, below (13), the square of $B_1 \cdots B_r$ is $(-1)^sI$, where $s = r(r+1)/2$. Hence $k^2 = 1$. *Thus the 2^{n-1} matrices (13) are linearly independent if $n \equiv 2$ (mod 4); while for $n \equiv 0$ (mod 4) they are either linearly independent or are connected by the relations which arise from $I = \pm B_1B_2 \cdots B_{n-1}$ by multiplication by*

the various matrices (13), *but are connected by no further irreducible linear relations.*

To illustrate this result, let $n = 4$. Then the 8 matrices

$$I, \quad B_1, \quad B_2, \quad B_3, \quad B_1B_2, \quad B_1B_3, \quad B_2B_3, \quad B_1B_2B_3$$

are either linearly independent or are connected by only four irreducible linear relations;

$$I = \pm B_1B_2B_3, \quad B_1 = \mp B_2B_3, \quad B_2 = \pm B_1B_3, \quad B_3 = \mp B_1B_2.$$

The latter express $B_1B_2B_3$, B_2B_3, B_1B_3, B_1B_2 linearly in terms of I, B_1, B_2, B_3, which are therefore in all cases linearly independent.

For any n, one of the reduced products of I and $B_1 \cdots B_{n-1}$ by any matrix (13) evidently contains fewer than half of the B's and the other contains more than half of the B's. Hence if irreducible linear relations exist, they serve merely to express the latter products in terms of the former. Thus in every case, the 2^{n-2} matrices (13) which are products of at most $(n-2)/2$ factors B are linearly independent.

But if we are given any $n^2 + 1$ matrices $(a_{ij}^{(k)})$ each with n rows and n columns, we can find numbers x_k not all zero such that

$$\sum_{k=1}^{n^2+1} x_k(a_{ij}^{(k)}) = 0,$$

i. e.,

$$\sum_{k=1}^{n^2+1} x_k a_{ij}^{(k)} = 0 \qquad (i, j = 1, \cdots, n),$$

since n^2 linear homogeneous equations in $n^2 + 1$ unknowns x_k have solutions not all zero.

Hence $2^{n-2} \leqq n^2$. This is satisfied if $n \leqq 8$, but fails if $n = 10$. But if it fails for $n = m$, it fails for $n = m + 1$, since

$$2^{m+1-2} = 2 \cdot 2^{m-2} > 2m^2 > (m+1)^2$$

if $(m-1)^2 > 2$, and hence if $m \geqq 3$. We have now proved that $n \leqq 8$.

The case $n = 6$ is readily excluded. Then the 2^5 matrices (13) are linearly independent. But $5 + 10 + 1$ of them are skew-symmetric (those with 1, 2 or 5 factors B). Between any 16 skew-symmetric six-rowed square matrices there exists a linear relation:

$$\sum_{k=1}^{16} x_k(b_{ij}^{(k)}) = 0; \qquad \sum_{k=1}^{16} x_k b_{ij}^{(k)} = 0 \quad (i, j = 1, \cdots, 6;\ i < j),$$

it being now necessary to examine only the 15 terms to the right of the main diagonal. But 15 linear homogeneous equations in 16 unknowns x_k have solutions not all zero.

THEOREM. *Except for $n = 1, 2, 4, 8$, there exists no identity* (8) *expressing the product $(x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2)$ as a sum of the squares of n bilinear functions of $x_1, \cdots, x_n$ and $y_1, \cdots, y_n$.*

HISTORY OF THE SUBJECT.

6. Gauss* remarked that the four square formula (4) is expressed in a simple way by

$$(Nl + Nm)(N\lambda + N\mu) = N(l\lambda + m\mu) + N(l\mu' - m\lambda'),$$

where l, m, λ, μ, λ', μ' are complex numbers, λ' being conjugate to λ, and μ' to μ, while Nl denotes the norm of l (§2).

7. C. F. Degen† extended Euler's formula (4) to eight squares:

$$\begin{aligned}
&(P^2 + Q^2 + R^2 + S^2 + T^2 + U^2 + V^2 + X^2) \\
&\qquad\qquad \times (p^2 + q^2 + r^2 + s^2 + t^2 + u^2 + v^2 + x^2) \\
&= (Pp + Qq + Rr + Ss + Tt + Uu + Vv + Xx)^2 \\
&+ (Pq - Qp + Rs - Sr + Tu - Ut + Vx - Xv)^2 \\
&+ (Pr - Qs - Rp + Sq \mp Tv \pm Ux \pm Vt \mp Xu)^2 \\
&+ (Ps + Qr - Rq - Sp \pm Tx \pm Uv \mp Vu \mp Xt)^2 \\
&+ (Pt - Qu \pm Rv \mp Sx - Tp + Uq \mp Vr \pm Xs)^2 \\
&+ (Pu + Qt \mp Rx \mp Sv - Tq - Up \pm Vs \pm Xr)^2 \\
&+ (Pv - Qx \mp Rt \pm Su \pm Tr \mp Us - Vp + Xq)^2 \\
&+ (Px + Qv \pm Ru \pm St \mp Ts \mp Ur - Vq - Xp)^2.
\end{aligned}$$

He stated [erroneously as we saw in §5] that there is a like formula for 2^n squares. For the case of 16 squares he gave the literal parts of the 16 bilinear functions, but left most of the signs undetermined, saying that the only difficulty is the prolixity of the ambiguities of signs. This paper has been overlooked by all subsequent writers on the subject.

8. J. T. Graves‡ communicated to W. R. Hamilton Jan. 18, 1844 (correcting some errors in signs in the formula communicated Dec. 26, 1843), a formula which differs from Cayley's (5) only in the interchange of 6 and 7, and a second formula which becomes Cayley's on writing $x_0, \cdots, x_7$ for $a, b, \cdots, h$. Hence Graves's formulas need not be inserted

* Posthumous MS., Werke, 3, 1876, 383–4.

† Mém. Acad. Sc. St. Pétersbourg, 8, années 1817–8 (1822), 207–219. There is a misprint in the sign of his term $\pm Rt$, here corrected.

‡ Proc. Roy. Irish Acad., 3, 1845–7, 527–9; Trans. Roy. Irish Acad., 21, II, 1848, 338–341; Phil. Mag. London, (3), 26, 1845, 320.

here. At first he expected that it would be possible to give an extension to 2^n squares.

9. J. R. Young's* formula, with s, t, u, v, y, z, w, x replaced by $a_1, \cdots, a_8$, is

$$(14)\quad \left(\sum_{i=1}^{8} a_i^2\right)(\Sigma\alpha_i^2) = (\Sigma a_i\alpha_i)^2 + (12 + 34 + 56 + 78)^2 + (13 + 42 + 57 + 86)^2 + (41 + 32 + 58 + 67)^2 + (15 + 62 + 73 + 48)^2 + (16 + 25 + 38 + 47)^2 + (17 + 82 + 35 + 64)^2 + (18 + 27 + 63 + 54)^2,$$

where ij denotes $a_i\alpha_j - \alpha_i a_j$. It was admitted to be equivalent to Graves's formulas. Young† stated that a like formula holds for 2^n squares, but soon afterwards admitted that this is erroneous, saying that he was prepared to prove that the proposition does not hold beyond 8 squares.

Young‡ gave a long discussion to show that the extension to a sum S_{16} of 16 squares is impossible. He exhibited a special relation $S_{16}S_{16}' = S_{16}''$ in which the roots of 8 of the squares in S_{16} are proportional to the roots of 8 of the squares in S_{16}'. He§ noted that, for $k = 2, 4, 8$, a product of a sum of km squares by a sum of kn squares can be expressed as a sum of kmn squares.

10. Cayley|| investigated the possibility of a formula for 2^n squares by introducing $2^n - 1$ symbols $a_0, b_0, \cdots$, not assumed to be commutative, but such that $a_0^2 = b_0^2 = \cdots = -1$ and

$$b_0c_0 = \pm a_0 = -c_0b_0, \quad c_0a_0 = \pm b_0 = -a_0c_0, \quad a_0b_0 = \pm c_0 = -b_0a_0.$$

Denoting this set of six equations by $a_0b_0c_0 = \pm$, let also $a_0d_0c_0 = \pm$, etc., where the sign is not necessarily the same as before, while the system of triples contains each duad once and but once, and the signs are to be chosen at will. Then

$$(w + aa_0 + bb_0 + \cdots)(w_1 + a_1a_0 + b_1b_0 + \cdots) = w_2 + a_2a_0 + b_2b_0 + \cdots,$$

where $w_2, a_2, \cdots$ are linear and homogeneous in $w, a, \cdots$ and in $w_1, a_1, \cdots$. Assume (I) that if any two triples with a common element, $e_0a_0b_0$ and $e_0c_0d_0$, occur in the system, there occur also $f_0a_0c_0, f_0d_0b_0, g_0a_0d_0, g_0b_0c_0$;

* Proc. Roy. Irish Acad., 3, 1845–7, 526–7.

† Phil. Mag., London, (3), 30, 1847, 424–5; 31, 1847, 123.

‡ Trans. Roy. Irish Acad., 21, II, 1848, 311–338. Outline in Proc. Roy. Irish Acad., 4, 1847–50, 19–20.

§ Phil. Mag., London, (3), 34, 1849, 114.

|| Phil. Mag., London, (4), 4, 1852, 515–9; Coll. Math. Papers, II, 49–52.

(II) that for any two pairs of triples, such as $e_0a_0b_0$, $e_0c_0d_0$ and $f_0a_0c_0$, $f_0d_0b_0$, the products of the signs of the triples in the first pair is the same as that in the second pair. Then

$$(w^2 + a^2 + b^2 + \cdots)(w_1^2 + a_1^2 + b_1^2 + \cdots) = w_2^2 + a_2^2 + b_2^2 + \cdots.$$

The converse was not proved, but it was stated that conditions (I) and (II) afford a complete test for the possibility of the 2^n square theorem.

T. P. Kirkman,* to whom Cayley had communicated privately the preceding test, verified that, for 15 elements $a, b, \cdots$, triples can be chosen so that (I) is satisfied, but that (II) then involves a contradiction.

11. F. Brioschi† showed that, if n is even, the square of the determinant $A = | a_{ij} |$ of order n is a skew-symmetric determinant $L = | l_{ij} |$ of order n with the general element

$$l_{rs} = a_{r1}a_{s2} - a_{r2}a_{s1} + a_{r3}a_{s4} - a_{r4}a_{s3} + \cdots + a_{rn-1}a_{sn} - a_{rn}a_{sn-1} = - l_{sr}.$$

Similarly, the square of $C = | c_{ij} |$ is $| p_{ij} |$, where $p_{rs} = c_{r1}c_{s2} - \cdots$. Let

$$AC = | A_{ij} |, \quad A_{rs} = \sum_{j=1}^{n} a_{rj}c_{sj}, \quad | A_{ij} |^2 = | L_{ij} |, \quad L_{rs} = A_{r1}A_{s2} - \cdots.$$

If the a's and c's are such that

$$(15) \qquad l_{12} = l_{34} = \cdots = l_{n-1n} = t, \quad p_{12} = \cdots = p_{n-1n} = u,$$

while the remaining l_{ij} and p_{ij} are zero, it is proved that $L_{12} = L_{34} = \cdots = L_{n-1n} = tu$, and that the remaining L_{ij} are zero. Now let $n = 8$ and take $a_{ii} = a_{11}$, $a_{ij} = - a_{ji}$ $(i \neq j)$ except for $a_{15} = a_{51} = a_{26} = a_{62} = a_{37} = a_{73} = a_{48} = a_{84}$ and take also

$$a_{12} = a_{43} = a_{56} = a_{87}, \quad a_{13} = a_{24} = a_{57} = a_{68}, \quad a_{14} = a_{32} = a_{58} = a_{76},$$

$$a_{16} = a_{47} = a_{52} = a_{83}, \quad a_{17} = a_{28} = a_{53} = a_{64}, \quad a_{18} = a_{36} = a_{54} = a_{72}.$$

Assume like relations between the c_{ij}. It is stated erroneously that relations (15) and the analogous relations between the A_{ij} hold, so that

$$\Sigma A_{1j}^2 = tu, \quad t = \Sigma a_{1j}^2, \quad u = \Sigma c_{1j}^2 \quad (j = 1, \cdots, 8).$$

Although $l_{12} = l_{34} = l_{56} = l_{78} = \Sigma a_{1j}^2$, it was pointed out by E. Sadun‡ that

$$l_{16} = 2(a_{11}a_{15} + a_{12}a_{16} + a_{13}a_{17} + a_{14}a_{18}) \neq 0,$$

so that we cannot make $t \equiv \Sigma a_{1j}^2$. In a footnote, Sadun reconstructed Brioschi's proof, and obtained (14) with 5 and 7, 6 and 8 interchanged.

* Phil. Mag., London, (3), 33, 1848, 447–459, 494–509; (3), 37, 1850, 292–301.

† Jour. für Math., 52, 1856, 133–141; Opere Mat., V, p. 511.

‡ Periodico di Mat., 14, 1899, 125–139; and pamphlet of 1896.

12. A. Lebesgue* gave an 8 square formula, communicated to him to Prouhet, which apart from signs becomes Cayley's formula (5) if we write $x_0, \cdots, x_7$ for $a, b, \cdots, h$.

13. A. Genocchi† concluded that sums of 2^n squares repeat under multiplication by an erroneous argument (false even for $n = 2$) based upon sums of two squares. The error was pointed out by Sadun (§11) and earlier by A. Puchta,‡ who interpreted the correct 8 square formula by means of regular bodies with 9 vertices in space of 8 dimensions.

14. E. Mathieu§ expressed Euler's identity (4) in the form

$$S_4 S_4' = S_4'', \quad S_4 = x_0^2 + x_1^2 + x_w^2 + x_{1+w}^2, \quad w^2 + w + 1 \equiv 0 \pmod 2,$$

$$x_0'' = x_0 x_0' + x_1 x_1' + x_w x_w' + x_{1+w} x_{1+w}', \qquad x_1'' = x_0 x_1' - x_1 x_0' - x_w x_{1+w}' + x_{1+w} x_w',$$

while x_w'' and x_{1+w}'' are derived from x_1'' by the substitution (z, wz), viz., $(1, w, 1 + w)$, on the subscripts. But S_4'' is unaltered also by $(0, w, 1 + w)$. Hence of the 24 permutations on the four subscripts, 12 give one decomposition into 4 squares, and 12 give another.

For $w^3 + w + 1 \equiv 0 \pmod 2$, Cayley's formula (5) can be expressed in the form

$$S_8 S_8' = S_8'', \quad S_8 = \Sigma x_j^2, \quad x_0'' = \Sigma x_j x_j',$$

$$x_1'' = x_0 x_1' - x_1 x_0' - x_w x_{1+w}' + x_{1+w} x_w' - x_{w^2} x_{1+w^2}'$$
$$+ x_{1+w^2} x_{w^2}' - x_{w+w^2} x_{1+w+w^2}' + x_{1+w+w^2} x_{w+w^2}',$$

where j ranges over the eight values $0, \cdots, 1 + w^2$ appearing in

$$s = (0)(1, w, w^2, 1 + w, w + w^2, 1 + w + w^2, 1 + w^2).$$

The remaining x_j'' are derived from x_1'' by applying this substitution s, which may be written in the form (w^z, w^{z+1}), the signs of the terms of x_j'' being determined so that the terms occurring in the above S_4'' occur with the same signs in S_8''. Now x_1'' is unaltered by (w^z, w^{2z}), while $x_0''^2, \cdots, x_7''^2$ are permuted by $(w^z, w^{1/z})$, where $1/z$ is replaced by the integer congruent to it modulo 7. Hence any symmetric function of these 8 squares is unaltered by the $3 \cdot 7 \cdot 8$ substitutions

$$(w^z, w^{z'}), \quad z' = \frac{az + b}{cz + d}, \qquad ad - bc \equiv 1, 2, 4 \pmod 7.$$

It is stated that these results cannot be extended to more than 8 squares.

* Exercices d'analyse numérique, 1859, 104; Introduction à la théorie des nombres, 1862, 65.

† Annali di Mat., 3, 1860, 202–5; Giornale di Mat., 2, 1864, 47–48.

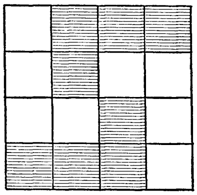
‡ Sitzungsber. Ak. Wiss. Wien (Math.), 96, II, 1887, 110–133.

§ Jour. für Math., 60, 1862, 351–6.

15. J. J. Thomson* verified Young's formula (14) by means of relations like

$$12\cdot 34 + 13\cdot 42 + 41\cdot 32 \equiv 0. \tag{16}$$

16. E. Lucas stated† that there is a relation between the formula expressing the product of two sums of n squares as a sum of n squares for $n = 4, 8, 16$, etc., and Sylvester's‡ square diagram formed of an equal number of white cases and black cases, such that for any two lines or two columns the number of variations of colors is always equal to the number of permanences. If, in the accompanying diagram, we replace each white case by a plus sign and each black case by a minus sign, we are led to Euler's formula (4).



17. S. Roberts§ argued that a 16 square formula is impossible. He assumed in effect that an m square formula must be of the type

$$\left(\sum_{i=1}^{m} a_i^2\right)\left(\sum_{i=1}^{m} c_i^2\right) = \sum_{i=1}^{m} (a_1 c_{i1} + \cdots + a_m c_{im})^2, \tag{17}$$

where $c_{i1}, \cdots, c_{im}$ and $c_{1i}, \cdots, c_{mi}$ are permutations of $\pm c_1, \cdots, \pm c_m$, and that the formula reduces to a $\frac{1}{2}m$ square formula by setting $a_i = c_i = 0$ $(i > m/2)$. In building the m square formula from the $\frac{1}{2}m$ square formula, he made free choice between the letters not already in the scheme. He derived an unique formula for $m = 4$ and for $m = 8$, but found after a tedious examination a contradiction for $m = 16$.

18. Cayley|| considered the linear algebra with the units $E_0 = 1, E_1, \cdots, E_7$, where $E_i^2 = -1$ $(i = 1, \cdots, 7)$ and

$$E_1E_2E_3 = \epsilon_1, \quad E_1E_4E_5 = \epsilon_2, \quad E_2E_4E_6 = \epsilon_4, \quad E_3E_4E_7 = \epsilon_6,$$

$$E_1E_6E_7 = \epsilon_3, \quad E_2E_5E_7 = \epsilon_5, \quad E_3E_5E_6 = \epsilon_7,$$

* Messenger Math., 7, 1877–8, 73–74.

† Assoc. franç. av. sc., 6, 1877, 213–4.

‡ Math. Quest. Educ. Times, 10, 1868, 74–6, 112 (diagrams for 8×8 squares and 16×16 squares). Cf. M. Jenkins, *ibid.*, 14, 1871, 22–25.

§ Quar. Jour. Math., 16, 1879, 159–170.

|| Amer. Jour. Math., 4, 1881, 293–6; Coll. Math. Papers, XI, 368–371. Incomplete summary in Johns Hopkins University Circulars, 1882, 203.

each ϵ_i being 1 or -1, and the first symbol denotes the six equations

$$E_1E_2 = \epsilon_1 E_3 = -E_2E_1, \quad E_2E_3 = \epsilon_1 E_1 = -E_3E_2, \quad E_3E_1 = \epsilon_1 E_2 = -E_1E_3.$$

For no values of the ϵ's is the algebra associative. We may set

$$(\Sigma a_i E_i)(\Sigma a_i' E_i) = \Sigma a_i'' E_i \quad (i = 0, 1, \cdots, 7).$$

Without loss of generality we may take $\epsilon_1 = \epsilon_2 = \epsilon_3 = +1$. Then

$$(\Sigma a_i^2)(\Sigma a_i'^2) \equiv \Sigma a_i''^2,$$

if and only if $-\epsilon_4 = \epsilon_5 = \epsilon_6 = \epsilon_7$. In such an algebra, $E_1 E_2 \cdot E_3 = E_1 \cdot E_2 E_3$ and similarly for each of the seven triads above. For the remaining 28 triads, $E_iE_j \cdot E_k = -E_i \cdot E_jE_k$. [We pass from one of these two algebras to the other by changing the signs of $E_2, \cdots, E_7$. If we take $\epsilon_4 = -1$ and change the sign of E_7, we get Cayley's earlier algebra (§4).]

19. Cayley* remarked that (16) establishes Euler's identity†

$$(\Sigma x_i^2)(\Sigma y_i^2) - (\Sigma x_i y_i)^2 = (12 + 34)^2 + (13 - 24)^2 + (14 + 23)^2.$$

The first step in forming this identity is to arrange the duads into a synthematic form: 12·34, 13·24, 14·23. The next step is to determine the signs. For 8 elements there is a single such synthematic arrangement; if 34, 56, 78 and each $1j$ are taken with positive signs, only one sign remains arbitrary, so that there are only two final schemes. For 16 elements, we have first to form 15 lines each containing the numbers $1, \cdots, 16$ in 8 duads, no duad being repeated. Only four types are found; for each it is found to be impossible to choose the signs. Cayley states that earlier writers had tacitly assumed that only one of the four types is possible and hence had not given a complete proof of the non-existence of the 16 square theorem. The question of the distinctness of the four types, apart from notation, was mentioned, but not discussed, by Cayley.

20. S. Roberts‡ remarked that Cayley's four types are all equivalent. But his directions for deriving the first from the second type are incorrect. Besides interchanging 13 with 14, and 15 with 16, and interchanging columns 13 and 14, and rows 15 and 16, it is necessary to interchange also rows 13 and 14, and columns 15 and 16. He indicated how his own process can be used to produce the four (equivalent) types.

21. F. Studnicka§ employed the product of two determinants:

$$\begin{vmatrix} a & b \\ -b' & a' \end{vmatrix} \cdot \begin{vmatrix} x & y \\ -y' & x' \end{vmatrix} = \begin{vmatrix} ax + by & -ay' + bx' \\ a'y - b'x & a'x' + b'y' \end{vmatrix}.$$

* Quar. Jour. Math., 17, 1881, 258–276; Coll. Math. Papers, XI, 294–313.

† Quoted in Math. Quest. Educ. Times, 75, 1901, 40.

‡ Quar. Jour. Math., 17, 1881, 276–280.

§ Sitzungsberichte K. Böhm Gesell. Wiss. Prag, 1883, 475–481.

Taking $a' = a, \cdots, y' = y$, we get (1). Next, let a' be the conjugate to the complex number $a, \cdots, y'$ the conjugate to y; we get Euler's (4). But he erred in employing the same formula when a' and a are conjugate quaternions, $\cdots, y'$ and y conjugate quaternions, to deduce the 8 square theorem, since he overlooked the fact that the initial formula holds only when multiplication is commutative [Vahlen, §27].

22. X. Antomari* wrote $(\alpha_i\beta_j)$ for $\alpha_i\beta_j - \beta_i\alpha_j$ and employed the identity

$$D = (\Sigma a_i x_i)(\Sigma b_i y_i) - (\Sigma a_i y_i)(\Sigma b_i x_i) = \Sigma(a_i b_j)(x_i y_j) \quad (i, j = 1, \cdots, 4; j > i).$$

In view of (16), written in a, b and again in x, y, we get

$$D = \{(a_1b_2) + (x_3y_4)\}\{(x_1y_2) + (a_3b_4)\}$$
$$+ \{(a_2b_3) + (x_1y_4)\}\{(x_2y_3) + (a_1b_4)\} + \{(a_1b_3) + (x_4y_2)\}\{(x_1y_3) + (a_4b_2)\}.$$

Taking a_j and x_j to be conjugate complex numbers and also b_j and y_j, for $j = 1, \cdots, 4$, we get an 8 square formula.

23. E. Lucas† stated that the determinant of the 8 equations

$$ax + by + cz + dt + ep + fq + gr + hs = X, \cdots, -hx + \cdots + as = S$$

is the fourth power of $\Delta = a^2 + \cdots + h^2$. To solve the equations, multiply them by $a, \cdots, h$, taken with proper signs, and add. We get

$$\Delta x = aX - bY - cZ - dT - eP - fQ - gR - hS, \cdots,$$
$$\Delta s = hX + \cdots + aS.$$

Squaring these and adding, we get $\Sigma a^2 \cdot \Sigma x^2 = \Sigma X^2$.

24. G. Arnoux‡ argued the impossibility of a 2^n square formula for $n > 3$.

25. Teilhet, de Montessus and Boutin§ gave special numerical examples of $S_r S_r' = S_r''$ for $r = 16$ and $r = 32$, where S_r is a sum of r squares.

26. E. Sadun‖ discussed m square formulas of the type (17). Since the product terms in $a_p a_t$ must cancel if $p \neq t$, we have

$$(18) \qquad c_{rp}c_{rt} = -c_{sp}c_{st}, \qquad c_{rp} = \pm c_{st}, \qquad c_{rt} = \mp c_{sp}.$$

Without loss of generality we may assume that the first row and first column of the matrix (c_{ij}) is $c_1, \cdots, c_m$. Then by (18) each diagonal term is $\pm c_1$, whence $m \neq 3$. In the rth and sth rows, $\pm c_{st}$ lies above c_{sp},

* Comptes Rendus, Paris, 104, 1887, 566–7.

† Théorie des nombres, 1891, 294.

‡ Assoc. franç. av. sc., 1896.

§ L'intermédiaire des math., 3, 1896, 259–262.

‖ Periodico di Mat., 14, 1899, 125–139; also as a pamphlet of 1896.

and $\mp c_{sp}$ above c_{st}. Hence m must be even. It is assumed that, if $a_i = c_i = 0$ $(i > m/2)$, (17) reduces to a $\frac{1}{2}m$ square formula. Thus if $m = 2^k\omega$, where ω is odd, we would get an ω square formula by continued halving. Hence $\omega = 1$, $m = 2^k$. The impossibility of a 16 square formula is established more simply than in the earlier papers.

27. K. Th. Vahlen* noted the error in Studnička's deduction of the 8 square theorem from the product of two two-rowed determinants and deduced that theorem by use of the product of two three-dimensional determinants [as had Antomari, §22]:

$$\begin{aligned}(aa' + bb' + cc' + dd')(xx' + yy' + zz' + tt') \\ &= (ax + by + cz + dt)(a'x' + b'y' + c'z' + d't') \\ &+ (-b'x + a'y + dz' - ct')(-bx' + ay' + d'z - c't) \\ &+ (-c'x - dy' + a'z + bt')(-cx' - d'y + az' + b't) \\ &+ (-d'x + cy' - bz' + a't)(-dx' + c'y - b'z + at').\end{aligned}$$

For $a = a'$, etc., this gives the formula (4) for 4 squares. If a' is the conjugate to a, it gives the 8 square theorem. He gave an analogous, much longer, formula which for $a = a'$, etc., becomes the 8 square formula, but when a' is the conjugate to a, etc., does not yield a 16 square formula.

28. E. Barbette† discussed the 4 and 8 square theorems in connection with magic squares.

* Giornale di Mat., 39, 1901, 181–4.

† Les sommes de p-ièmes puissances . . . , Liège, 1910.